

In the claims:

Please substitute the following full listing of claims for the claims as originally filed or most recently amended.

1. (Currently Amended) A method for resource allocation where only the utility is not evaluated but a marginal utility is ~~known~~ evaluated, comprising the steps of:

determining an initial step size;

evaluating step size effectiveness with only information on the marginal utility and without evaluating utility;

reducing the step size if necessary based on said marginal utility; and

deploying an effective step size for utility optimization.

2. (Original) The method of claim 1, wherein allocated resources are database memory components.

3. (Original) The method of claim 1, wherein allocated resources are computer memory resources.

4. (Currently Amended) The method of claim 1, wherein allocated resources are computer system resources, and

resources are allocated to achieve optimal service level objectives.

5. (Currently Amended) An apparatus for resource allocation comprising:

- a step size determination engine;
- an optimization module for evaluating step size effectiveness with ~~only~~ information on the marginal utility but without evaluation of utility;
- a constraint module for reducing the step size if necessary based on said marginal utility; and
- an interface module for deploying an effective step size for utility optimization.

6. (Original) The apparatus of claim 5, wherein allocated resources are database memory components.

7. (Original) The apparatus of claim 5, wherein allocated resources are computer memory resources.

8. (Original) The apparatus of claim 5, wherein allocated resources are computer system resources and resources are allocated to achieve optimal service level objectives.

9. (Currently Amended) A method of maximizing or minimizing an objective function $f(x)$, subject to constraints on a vector x where each vector x denotes a particular allocation of resources and the constraints generally describe properties of the resources which must be satisfied, the method maximizing or minimizing the objective function $f(x)$ while satisfying the constraints on x without use of a knowledge of f , said method comprising the steps of:

- starting from an initial allocation, calculating a marginal utility of said allocation;
- calculating constraint functions of said allocation;
- applying the calculated constraint function information and marginal utility information to obtain a next allocation;
- repeating the steps of calculating a marginal utility without evaluation of f , if known, calculating constraint functions and applying the calculated constraint function information and marginal utility information until a stopping criteria is satisfied; and
- returning a locally optimal allocation of resources.

10. (Original) The method of maximizing or minimizing an objective function $f(x)$ as recited in claim 9, wherein said marginal utility is the gradient of the function f .

11. (Original) The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 9, wherein only the gradient ∇f is known and there is no procedure to evaluate the objective function f .

12. (Original) The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 9, wherein only the gradient ∇f and the Hessian $\nabla^2 f$ are known and there is no procedure to evaluate the objective function f .

13. (Original) The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 9, wherein the objective function is a utility function and the method maximizes the utility function.

14. (Original) The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 13, wherein said utility function is time saved.

15. (Original) The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 13, wherein said utility function is utilization of computer processors.

16. (Original) The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 13, wherein said utility function is a number of transactions processed.

17. (Original) The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 13, wherein said utility function is utilization of computer memory.

18. (Original) The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 9, wherein the objective function is a cost function and the method minimizes the cost function.

19. (Original) The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 18, wherein said cost function is power consumption.

20. (Original) The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 18, wherein said cost function is total disk input/output time.

21. (Original) The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 18, wherein said cost function is average system response time.

22. (Original) The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 9, wherein a backtracking line search is implemented in which a step α is decreased until the following condition is satisfied:

$$(\nabla f(x+\alpha p) - c_1 \nabla f(x))^T p \leq 0,$$

where p is a search direction, c_1 is a constant, and

$(\nabla f(x+\alpha p) - c_1 \nabla f(x))^T$ is the transpose of $(\nabla f(x+\alpha p) - c_1 \nabla f(x))$.

23. (Original) The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 9, the method optimally allocating resources wherein the allocated resources are database memory components.

24. (Original) The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 9, the method optimally allocating resources wherein allocated resources are computer memory resources.

25. (Currently Amended) The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 9, the method optimally allocating resources wherein

allocated resources are computer system resources, and

resources are allocated to achieve optimal service level objectives